

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 21:32:16 ; Search time 43 Seconds
(without alignments)
631.913 Million cell updates/sec

Title: US-09-619-032a-4

Perfect score: 1879

Sequence: 1 LRAIVFHGNLQYAEIPKSEI.....RLDAFRAIYNDMGNGEPP 364

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1879	100.0	364	US-10-166-606-4	Sequence 4, Appl 1
2	1589	84.6	346	US-08-613-220B-4	Sequence 4, Appl 1
3	140.5	7.5	647	US-07-894-212A-8	Sequence 8, Appl 1
4	140.5	7.5	649	US-07-894-212A-2	Sequence 2, Appl 1
5	140.5	7.5	650	US-07-893-928A-1	Sequence 1, Appl 1
6	105	5.6	227	US-09-107-532A-5554	Sequence 5554, Ap
7	101.5	5.4	360	US-09-710-279-2150	Sequence 2150, Ap
8	101.5	5.4	376	US-09-710-279-2106	Sequence 2106, Ap
9	101	5.4	329	US-08-270-013B-2	Sequence 2, Appl 1
10	101	5.4	329	US-08-838-418-2	Sequence 2, Appl 1
11	97	5.2	653	US-09-540-236-3128	Sequence 3128, Ap
12	94.5	5.0	367	US-09-134-001C-3678	Sequence 3678, Ap
13	92	4.9	375	US-09-710-279-1840	Sequence 1840, Ap
14	92	4.9	845	US-09-248-796A-15394	Sequence 15394, A
15	91.5	4.9	741	US-09-252-991A-31448	Sequence 31448, A
16	91.5	4.9	1009	US-09-693-146-4	Sequence 4, Appl 1
17	91	4.8	1031	US-09-543-681A-8245	Sequence 8245, Ap
18	90.5	4.8	406	US-09-134-001C-3570	Sequence 3570, Ap
19	90.5	4.8	789	US-09-248-796A-19294	Sequence 19294, A
20	90.5	4.8	1171	US-09-248-796A-16043	Sequence 16043, A
21	90	4.8	3854	US-09-949-016-7876	Sequence 7876, Ap
22	89.5	4.8	322	US-09-489-039A-12383	Sequence 12383, A
23	89	4.7	857	US-09-248-796A-20522	Sequence 20522, A
24	88.5	4.7	399	US-09-252-991A-23741	Sequence 23741, A
25	88.5	4.7	764	US-09-235-451-36	Sequence 36, Appl 1
26	88.5	4.7	764	US-09-978-303-36	Sequence 36, Appl 1
27	88.5	4.7	993	US-09-894-998A-50	Sequence 50, Appl 1

28	88.5	4.7	993	US-10-237-551-50	Sequence 50, Appl 1
29	88.5	4.7	1037	US-09-894-998A-54	Sequence 54, Appl 1
30	88.5	4.7	1037	US-10-237-551-54	Sequence 54, Appl 1
31	88.5	4.7	1113	US-09-894-998A-51	Sequence 51, Appl 1
32	88.5	4.7	1113	US-10-237-551-51	Sequence 51, Appl 1
33	88.5	4.7	1114	US-10-237-551-202	Sequence 202, Appl 1
34	88.5	4.7	2089	US-08-418-893D-23	Sequence 23, Appl 1
35	88.5	4.7	2089	US-08-418-893D-24	Sequence 24, Appl 1
36	88	4.7	329	US-09-710-279-1460	Sequence 1460, Ap
37	88	4.7	339	US-09-328-352-4144	Sequence 4144, Ap
38	88	4.7	339	US-09-489-039A-11436	Sequence 11436, A
39	88	4.7	420	US-09-248-796A-23483	Sequence 23483, A
40	88	4.7	1151	US-09-710-279-2448	Sequence 2448, Ap
41	88	4.7	1154	US-09-134-001C-3428	Sequence 3428, Ap
42	87.5	4.7	542	US-09-949-016-6778	Sequence 6778, Ap
43	87.5	4.7	554	US-09-949-016-11687	Sequence 11687, A
44	87.5	4.7	689	US-08-248-021A-2	Sequence 2, Appl 1
45	87	4.6	235	US-09-141-135-2	Sequence 2, Appl 1

ALIGNMENTS

```
RESULT 1
US-10-166-606-4
; Sequence 4, Application US/10166606
; Patent No. 6644756
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; APPLICANT: Reid, John
; TITLE OF INVENTION: ALPHA GALACTOSIDASES AND METHODS FOR
; FILE REFERENCE: 09010-004005
; CURRENT APPLICATION NUMBER: US/10/166,606
; PRIOR FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 09/407,806
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-166-606-4

Query Match      100.0%  Score 1879; DB 4; Length 364;
Best Local Similarity 100.0%  Pred. No. 5.2e-191;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRAIVFHGNLQYAEIPKSEIPVIEIKAYIPVETLIKESIPGALITGYTLKPLKDIID 60
   1 LRAIVFHGNLQYAEIPKSEIPVIEIKAYIPVETLIKESIPGALITGYTLKPLKDIID 60
DB 1 LRAIVFHGNLQYAEIPKSEIPVIEIKAYIPVETLIKESIPGALITGYTLKPLKDIID 60
   1 LRAIVFHGNLQYAEIPKSEIPVIEIKAYIPVETLIKESIPGALITGYTLKPLKDIID 60
QY 61 LVKGIADLIRITSTYTHALPLPLSRVBAQVQREKSEELFEVSPKFWLPYLA 120
   61 LVKGIADLIRITSTYTHALPLPLSRVBAQVQREKSEELFEVSPKFWLPYLA 120
DB 61 LVKGIADLIRITSTYTHALPLPLSRVBAQVQREKSEELFEVSPKFWLPYLA 120
   61 LVKGIADLIRITSTYTHALPLPLSRVBAQVQREKSEELFEVSPKFWLPYLA 120
QY 121 DRIIPAIKNDGYELFPDGEAMLPFAHLSAIKIKPLPYLILAAOREKFRYSYLLG 180
   121 DRIIPAIKNDGYELFPDGEAMLPFAHLSAIKIKPLPYLILAAOREKFRYSYLLG 180
DB 121 DRIIPAIKNDGYELFPDGEAMLPFAHLSAIKIKPLPYLILAAOREKFRYSYLLG 180
   121 DRIIPAIKNDGYELFPDGEAMLPFAHLSAIKIKPLPYLILAAOREKFRYSYLLG 180
QY 181 LRELRKAIKLVFEGKVTLLKAVKDIEAVPVVAVNTAVNLGIGRLPLMNPCKVASWIEDK 240
   181 LRELRKAIKLVFEGKVTLLKAVKDIEAVPVVAVNTAVNLGIGRLPLMNPCKVASWIEDK 240
DB 181 LRELRKAIKLVFEGKVTLLKAVKDIEAVPVVAVNTAVNLGIGRLPLMNPCKVASWIEDK 240
   181 LRELRKAIKLVFEGKVTLLKAVKDIEAVPVVAVNTAVNLGIGRLPLMNPCKVASWIEDK 240
QY 241 NILVGTIERTIGRDINGRMSVGLLEVDLNSCLPSGLKSGRELYLTRSSWAP 300
   241 NILVGTIERTIGRDINGRMSVGLLEVDLNSCLPSGLKSGRELYLTRSSWAP 300
DB 241 NILVGTIERTIGRDINGRMSVGLLEVDLNSCLPSGLKSGRELYLTRSSWAP 300
   241 NILVGTIERTIGRDINGRMSVGLLEVDLNSCLPSGLKSGRELYLTRSSWAP 300
QY 301 DKSLIRFEDSGNRLNMLSYNMGELAFLENSDARGMBPLPERRLDAPRAIYNDMG 360
   301 DKSLIRFEDSGNRLNMLSYNMGELAFLENSDARGMBPLPERRLDAPRAIYNDMG 360
```

Db 301 DKSRLRWDEGNARLNLMSYNNRGELAFLAENSDARGWEPLPERRLDAPRAIYNDWRGE 360
Qy 361 NGEP 364
Db 361 NGEP 364

RESULT 2

US-08-613-220B-4
; Sequence 4, Application US/08613220B
; Patent No. 5958751
; GENERAL INFORMATION:
; APPLICANT: Murphy, Dennis
; APPLICANT: Reid, John
; TITLE OF INVENTION: ALPHA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/613,220B
; FILING DATE: 08-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hallie, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-68-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 346 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-613-220B-4

Query Match 84.6%; Score 1589; DB 2; Length 346;

Best Local Similarity 95.1%; Pred. No. 3.3e-160;

Matches 346; Conservative 0; Mismatches 0; Indels 18; Gaps 18;

Qy 1 LRALVFHGNQVAYEIPKSEIPKVIETLIKESIPFGNITGYTLKFLPKXIID 60
Db 1 LRLALVFHGNQVAYEIPKSEIPKVIETLIKESIPFGNITGYTLKFLPKXIID 57
Qy 61 LVKGIASDLIEIIGTSYTHAIIPLPLSRVEAQVORDREVKEELFEVSPKGFWLPETAY 120
Db 58 LVKGIASDLIEIIGTSYTHAIIPLPLSRVEAQVORDREVKEELFEVSPKGFWLPETAY 114
Qy 121 DPLIPALKNKGVEYLFADGAMLFSAHLSAIPKIPLYPHILKAQREKFRYISYLLG 180
Db 115 DPLIPALKNKGVEYLFADGAMLFSAHLSAIPKIPLYPHILKAQREKFRYISYLLG 171
Qy 181 LRELKRAIKLVFGSKYTLKAVKDIKAVVAVVAVNTAVMLGIGRLPLMNPKKVSWIEMDK 240
Db 172 LRELKRAIKLVFGSKYTLKAVKDIKAVVAVVAVNTAVMLGIGRLPLMNPKKVSWIEMDK 228
Qy 241 NULLGTIDIEFIGYDIAGYRMSVEGLLEVIIDELNSELCLPSELKHSGRELYLRTSSWAP 300

Db 229 NULLGTIDIEFIGYDIAG-RMSVEGLEVIIDELNSELCLPSELKHSGRELYLRTSSWAP- 285
Qy 301 DKSRLRWDEGNARLNLMSYNNRGELAFLAENSDARGWEPLPERRLDAPRAIYNDWRGE 360
Db 286 DKSRLRWDEGNARLNLMSYNNRGELAFLAENSDARGWEPLPERRLDAPRAIYNDWRG- 342

Qy 361 NGEP 364
Db 343 NGEP 346

RESULT 3

US-07-894-212A-8
; Sequence 8, Application US/07894212A
; Patent No. 5366883
; GENERAL INFORMATION:
; APPLICANT: ASADA, KIYOZO
; APPLICANT: UEMORI, TAKASHI
; APPLICANT: MURAI, HIROYUKI
; APPLICANT: KATO, IKUNOSHIN
; APPLICANT: LADERMAN, KENNETH
; APPLICANT: ANFINSEN, CHRISTIAN
; TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DABRY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/894,212A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 95469/C-1195
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 647 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-894-212A-8

Query Match 7.5%; Score 140.5; DB 1; Length 647;

Best Local Similarity 22.8%; Pred. No. 9.3e-06;

Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;

Qy 25 EKAYIPVETIEIKESIP---FGNITGYTLKFLP---KQIIDLVKGIASDLIEIIGTSY 78
Db 27 EKAYIPVETIEIKESIP---FGNITGYTLKFLP---KQIIDLVKGIASDLIEIIGTSY 78
Qy 79 THAIIPLPLSRVEAQVORDREVKE--ELFEVSPKGFWLPETAYDPLIPALKNKGVEYL 136
Db 85 YEPVLASIP---KEDRIEQRLMKWEKASIGFDARGVWLTQVPELVTKLSSGIDYV 141
Qy 137 FADGAMLFSAHLSAIPKIPLY-PHILKAQRE-----KFRYISYLLGRELKRA 187
Db 142 IVD-----DYHMSAGLSKEELWYWPYTDGGEVIAVFPIDELKLR---YLIFRPADV 192

QY 188 IRL---VFEGKVTLKAV--KDIEAVPVWVAVNTAVMLGIGRLPLMPPKVASWI----- 236
DB 193 LEYLSLIDGDESSKVAVFHDDEKFGIMGTETWY-----EKGMLREFDOR 239
QY 237 ---EDKNDILLYGTIDIE-----FIGYRDIAGYRMSVEGLLEVIIDELINSELCLP----- 281
DB 240 ISSDEKINMLYTEYLEKYPKPGVLVLPISY-----FEM-----SEMSLPKQARLP 287
QY 282 ---SELKHSG---RELTLRTSSWAPDPSLRIMREDEGN---ARLNLMLSYNMRGELATL 330
DB 288 VEFVVELKVKGIPEKRYRVRVGGIW---KNF-FYKYPBSNMYMKMLWVSKLVRNN----- 339
QY 331 AENSDARGWEPLPERRLDAFRAIYND--WRG 359
DB 340 -----PEARKYILRAQCNDAYWHG 358

RESULT 4
US-07-894-212A-2
Sequence 2, Application US/07894212A
Patent No. 536883
GENERAL INFORMATION:
APPLICANT: ASADA, KIYOZO
APPLICANT: UEMORI, TAKASHI
APPLICANT: MUKAI, HIROYUKI
APPLICANT: KATO, IKUNOSHIN
APPLICANT: LADERMAN, KENNETH
APPLICANT: ANFINSEN, CHRISTIAN
TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/894,212A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 95469/C-1195
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 649 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-894-212A-2

Query Match 7.5%; Score 140.5; DB 1; Length 649;
Best Local Similarity 22.8%; Pred. No. 9.4e-06;
Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;

QY 25 EKAYPIVETILKEIP---FGLNTGYTLKFLP---KDIIDLKVGASLIDLIETIGSY 78
DB 29 EKCYPFLETL--BEYPMNKVAIHITSGPLIEWLQNNRPYIDLNSLVKRGQVEIVAGF 86
QY 79 THAIPPLPLSRVNAQVORDREVEK--ELFEVSPKGFMLPELAYDPIIPALIKONGEYVL 136
DB 87 YEPVLASIP---KEORIRIQIRLMKEMAKSIGFDAGVWLTERVWQPELVKTLKESGIDIV 143

QY 137 PADGEAMLEPSAALNGAIKIPLY-PHLIKAQB-----KRFRYISYLLGLRELKA 187
DB 144 IVD-----DYHPSAGLSKEELWYPTTEDGEVIAVPEIDELR---YLLPFRVDRKY 194
QY 188 IRL---VFEGKVTLKAV--KDIEAVPVWVAVNTAVMLGIGRLPLMPPKVASWI----- 236
DB 195 LEYLSLIDGDESSKVAVFHDDEKFGIMGTETWY-----EKGMLREFDOR 241
QY 237 ---EDKNDILLYGTIDIE-----FIGYRDIAGYRMSVEGLLEVIIDELINSELCLP----- 281
DB 242 ISSDEKINMLYTEYLEKYPKPGVLVLPISY-----FEM-----SEMSLPKQARLP 289
QY 282 ---SELKHSG---RELTLRTSSWAPDPSLRIMREDEGN---ARLNLMLSYNMRGELATL 330
DB 290 VEFVVELKVKGIPEKRYRVRVGGIW---KNF-FYKYPBSNMYMKMLWVSKLVRNN----- 341
QY 331 AENSDARGWEPLPERRLDAFRAIYND--WRG 359
DB 342 -----PEARKYILRAQCNDAYWHG 360

RESULT 5
US-07-893-928A-1
Sequence 1, Application US/07893928A
Patent No. 5578479
GENERAL INFORMATION:
APPLICANT: LADERMAN, KENNETH
APPLICANT: ANFINSEN, CHRISTIAN
TITLE OF INVENTION: a-AMYLASE FROM HYPERTHERMOPHILIC
ARCHAEBACTERIUM
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Tape
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/893,928A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: 95470/C-1197
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-893-928A-1

Query Match 7.5%; Score 140.5; DB 1; Length 650;
Best Local Similarity 22.8%; Pred. No. 9.4e-06;
Matches 89; Conservative 60; Mismatches 127; Indels 115; Gaps 23;

QY 25 EKAYPIVETILKEIP---FGLNTGYTLKFLP---KDIIDLKVGASLIDLIETIGSY 78
DB 29 EKCYPFLETL--BEYPMNKVAIHITSGPLIEWLQNNRPYIDLNSLVKRGQVEIVAGF 86
QY 79 THAIPPLPLSRVNAQVORDREVEK--ELFEVSPKGFMLPELAYDPIIPALIKONGEYVL 136

```

Db      87 YEEVLASIP---KEDRIEQRRLMKEMAKSIGPDARGVWTERWQPELVYTKLESIDYV 143
Qy      137 FADGEAMLSAHLNSAIKPKPLY-PHLIAQRE-----KRFRTSYTLGLRELKRA 187
Db      144 IYV-----DYHFMSSAGLSKEBELYWPYTEDGEVLAVPIDELKLR---YLIFPRPVDKV 194
Qy      188 IKL---VFEGKVTLKAV--KDIEAVPVWAVNTAVMLGIGRLPLMNKVASNY-----236
Db      195 LEVLHSLIDDEKSAVAFHDDGEKFGIWPGETYEMV-----EKGMLREFDR 241
Qy      237 ---EDKNILLYGTDIE---FIGYDIAGYMSVEGLLEVIDELNSLCLP-----281
Db      242 ISSDEKINILMYTEYLEKPKYRGVLVLPILASY-----PEM-----SEWSLPAKQARLF 289
Qy      282 ----SELKHS---RELVRTSSMAPDKSLRTWREDEGN---ARLMLSYNMRGELATL 330
Db      290 VEFVNELKVAGIEKRYKRVFRGGIW---KNF-FYKYPESNYMKRMLMVSGLVNN---341
Qy      331 AENSDBGWEPLPERRLDAFRAIYND--WRG 359
Db      342 -----PEARKYLIRACQNDAYWHG 360

```

```

RESULT 6
US-09-107-532A-5554
; Sequence 5554, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 893-5007
; TELEFAX: (781) 893-8277
; INFORMATION FOR SEQ ID NO: 5554:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 227 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULAR TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...227
; SEQUENCE DESCRIPTION: SEQ ID NO: 5554:

```

```

US-09-107-532A-5554
Query Match      5.4%; Score 105; DB 4; Length 227;
Best Local Similarity 24.2%; Pred. No. 0.01;
Matches 60; Conservative 34; Mismatches 88; Indels 66; Gaps 12;
Qy      140 GEAMLSAHLNSAIKPKPLYPHLIAQREKRRYSYTLGLRELK-----AI 188
Db      3 GEOMISITMKLV-----FGLLGLLVALLGKKSSSEITPPDLVY 43
Qy      189 KLYPEGVTLKAVD-----IEAVPVWAVNTAVML-GIGRLPLMNPKVASWIEDKD 240
Db      44 TLVGLGILEESTYDDVWVGHVFAIALW---AWIYGERIYVOKN-EKVRWVWGP 97
Qy      241 NILLYGTDIEFGYRDIAGYMSVEGLLEVIDELNSLCLPSE-----LKHSGRELYR 294
Db      98 SVLIKD---GVIMTELTNHHEBQRAI---LRQECFFLENAKHVLLENAGQMSVLK 151
Qy      295 TSSMAPDKSLRTWREDEGNARLMLSYNMRGELAFLAENSDBGWEPLPERRLDAFRAIY 354
Db      152 KSD--EDKALSTILVDEGQIQHKVLSNQLTE-AWIMENLKKEGY-----ADVQQLIY 201
Qy      355 NDMRGENG 362
Db      202 VEWSESEKG 209

```

```

RESULT 7
US-09-710-279-2150
; Sequence 2150, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2150
; LENGTH: 360
; TYPE: JRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-09-710-379-2150

```

```

Query Match      5.4%; Score 101.5; DB 4; Length 360;
Best Local Similarity 19.1%; Pred. No. 0.051;
Matches 57; Conservative 52; Mismatches 117; Indels 73; Gaps 12;
Qy      4 LVFHGNLQVAIRPKSEIPKVIKRAYIPVETLKEIIPGLNITGYTLKFLPRDIIDLVK 63
Db      61 LIRKG-----VKRIVEDGYSIIRKLIQNNI--NLIALHTNLVDNPGVNRMLA 107
Qy      64 GGIASDLIELIGT--SYTHAIIPLPLSRVEAOVQDREYKELFEVSPGFWLPELAYD 121
Db      :.08 DQIGLENISMINNSSYYKVVQTFIPKNYIE-----DFKSLNLE-----147
Qy      :.22 PIRPILKONGEYELFADGEAMLSAHLNSAIKPKPLYPHLIAQREKRRYSYTLGL 181
Db      148 ----GLAKGNYEYCFPESEG-----KQGFEPVGDASFYIKLS-----IEV---187
Qy      182 RELKRAIKLVF-----EGVTLKAVKDIE--AVPV--WVAVNTAVMLGIGRLPLMNPKV 232
Db      188 ----DEIKLEFMKONLEIITKRAIILDNHYETPVPDFIOMNKESEYGLGIQGLNQTM 243
Qy      233 ASWIEDKNILLYGTDIEFGYRDIAGYMSVEGLLEV-----IDELNSLCLPSELKH 286
Db      244 LDEFSEYAKQNLIPSVRYTGGHDSPIKVAIIIGSGIGPEYKASQIGADVFTGDIKH 302

```


[illegible]

```

Db      344 VKTGEVLAMGSPENNNLSERDANER 372

RESULT 12
US-09-134-001C-3678
; Sequence 3678. Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO: 3678
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3678

Query Match          5.0%; Score 94.5; DB 3; Length 367;
Best local similarity 18.7%; Pred. No. 0.29;
Matches 56; Conservative 52; Mismatches 110; Indels 73; Gaps 12;

Cy      4 LVFNGNLQYAIRPSEIRPKVIEKAYIPVIETLKKEIRPGFNITGTTLKFLPKDIIDLVK 63
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      68 LIFFG-----VRIVEDGGSTIRKLQNNT--NLALHTNDLVNKGVNRMLA 114

Cy      64 GLIASDLIEIGT--SYTHAILPLPLSRVEAQVQRDRVEKSELFEVSPKGFWLPBLAYD 121
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      115 DQIGENISMINTSSYYKKQTFRPKRYI-----DFQDSLNEI----- 154

Cy      122 PIIPAALKONGEYFLADGEMLFSAHLNSAIKPPIPLYPHILIKAKREKRFYISYLGL 181
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      155 ---GLAKEGNVEYCFEFSEG-----KQGFGPVGDASPYIGKIDS-----IEVV--- 194

Cy      182 RELKATLVF----BEKVTLKAVKOLE--AVPP--WVAVTAAVMLGRLPLNNPKV 232
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      195 ---DEIKLEPMIKONEIEITRKAILDNHPYETPAFDRIKNKKESEYGGLIGQLNQWT 250

Cy      233 ASWIEDKNLLGYGDIFIGRIDYAGRMVGELLEV-----IDELNSELCIPSELKH 286
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      251 LDSESEYAKQUNIPSVARYTGQHDSPICKVALIGSGIGFEYKAQLGADVTVTDIGH 309

RESULT 13
US-09-710-279-1840
; Sequence 1840. Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUJ480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1840
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1840

```

Query Match 4.9%; Score 92; DB 4; Length 375;
Best Local Similarity 19.3%; Pred. No. 0.55;
Matches 72; Conservative 55; Mismatches 109; Indels 138; Gaps 16;

QY 59 IDLVKGIASDLIEITGTSYTHAIPLLPSRV-----EAQVORDREVEKELEFVSPK--- 111
DB 3 IDIEBG-----DPIAIGTSGSGKTALNMNMISTEGEITIDGKNIKELPVELRBSI 58
QY 112 GFMLPELAYDP-----IIPAILK-----DNGEYLT----- 136
DB 59 GYVIGQIGLMPMTVENIVLPKLLKWSQEKDEKAKELIRLVLPSEYLDREYSELG 118
QY 137 -----FADEANLFSANLNSAIKPI-KPIYPLIK--AQREKFRYISYLG 181
DB 119 GQOQRIQVRLAAEODIILWDEPFGLDPIRDLQVLKQLQOQLGTFIFVTH--- 174
QY 182 RELKAIKL-----VEGKVTLKAVD 203
DB 175 -DMDEATLADKICIMTNGQVITYPTDNILRSPANDFRDPTIGQNRLLIQDRPNIRTYKD 233
QY 204 IEAVPVVAVNTAVMLGIGRLPLMPKVASWIEDKNILLYGTIDIEFGYRDIAGYMS 263
DB 234 AMIKPVTVHVDRLNDAYN---IMREKRY-----DTIFVGNDEHLGYLIDIEDING 283
QY 264 VEGGLEVIDELNSELG---LSEELKSGRELYLRTSSMAP-----DKSL----- 304
DB 284 LRHHKELIDTWQRIQVRLRIDSKLQDSVRTLLKRVNRVNVVSDNKTLLGLVTRANLVD 343
QY 305 ----RIMRE-DEGN 313
DB 344 IVYDSIMGELESN 357

RESULT 14
US-09-248-796A-15394
Sequence 15394, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Kelch Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 15394
LENGTH: 845
TYPE: PRT
ORGANISM: Candida albicans
US-09-248-796A-15394

Query Match 4.9%; Score 92; DB 4; Length 845;
Best Local Similarity 20.2%; Pred. No. 2.1;
Matches 78; Conservative 64; Mismatches 117; Indels 128; Gaps 21;

QY 38 BEIP-FGLNTGYTLKPLPKDIIIDLVKGIASDLIEITGTSYTHAI 82
DB 202 DEMHITGIDING-----KRIKMPAKGSALDQLESIDLBEGTGLDONTGTS----- 249
QY 83 LPLPLPSRVAQVORDREVEKELEFV-SPKGFMLPELAYDPPIIPAILKONGEYLFADGE 141
DB 250 ---LKTDEBELRLKIQOQENTDENINP-----YEPIDIMFTD----- 286
QY 142 AMLESANLSAIKPIKPIYPLIKAQREKFRYISYLGRLKRAIKALYVEGKY----- 196
DB 287 -----EELIMP-----VTAVPKPKRFVPSKHEAKRVKIVAKIREGRILPPNK 329
QY 197 -----TLKAVKDIEAVPVW---VAVNTAVM-LGIGRLP-----LMPKPKVA 233

DB 330 VAKQULTEEBEEDQNFPLMDQDEITISDHIMLRAPKLPPTNBSYNPEEVLITBEES 389
QY 234 SWIE-----DKONILLYGTIDIEFGYRDIAGYMSVEGLEVIDELNSELCLPSELKHSGR 289
DB 390 KWLQSGPIDRENFL---PKYNSLRQVPGVDSVRERF-----EKSLDYLAPVRHN-- 440
QY 230 ELYLRTSSMAPD---KSLR-----IWEDEGNARLMLSYNKGELAFLENSD-- 335
DB 441 KLNIDPDSILPDLPSPKDLRPPRIRCSYIEGHTG--KIRTSIDPOG--LWLATGSDG 496
QY 336 -ARGMEPLPERLDAFRAIYNDMRGEN 361
DB 497 SVRIWEILTGROVYKIQLINKENINED 523

RESULT 15
US-09-252-991A-31448
Sequence 31448, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31448
LENGTH: 741
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31448

Query Match 4.9%; Score 91.5; DB 4; Length 741;
Best Local Similarity 19.5%; Pred. No. 1.9;
Matches 84; Conservative 63; Mismatches 131; Indels 153; Gaps 22;

QY 34 TLIEEIPFGLNTGYTLKPLPKDIIIDLVKGIASDLIEITGTSYTHAI-----LPLLP 87
DB 257 TMMKISHPI---VFCHAVSYVKDFD---KWG---QLFEEELGVNPNNGISSVYDKIKSLP 308
QY 88 LSRVQAQVORDREVEKELEFV-----SPKG---FWLP-ELAVDPITPAIKDNG 132
DB 309 ASQGE-----ELIHDHEVYSHRPEAMVDVKGITNMLHPSVYIDASMPAIRNSG 361
QY 133 YEYLFADG-----EAMLFSAHLSAIKPIKPIYPLIKAQREKFRY 174
DB 362 -QMWGDKQKQDTKAVMPESTYARIYQEMINFCNTNGAFDPTMGSVNPGIMAQAEY 420
QY 175 ISYLLGELRLKAIKLVPEGVTLKAV-----KQIEAVPVVAVNTA-----VM 218
DB 421 GSH-----DKTEEMTADG---TWKVLVADSSVLMQHDVETGDIWPAQOTDAPIRDWK 471
QY 219 LGIGR-----LPLMP-----KQVASWIEDKNILLYGTIDIEFGYRDIAGY 261
DB 472 LAVTRARQSDPRLPFWLDPBAHRELKRYELVYKHD---LIGLDISIMQYN--AIR 526
QY 262 MSVEGLEVIDELN-----SEL-----CL 280
DB 527 VSMERLLKQDTISVGNVLRLDYLTDLPPIMELGTSANLIVPLMAGGWETGAGGSA 586
QY 281 PSELKSGRELYLRTSSAPPKSLRIMED---EGNARLMLSYNKGELAFLENSDA 336
DB 587 PKHVOQLVEENYLLMDSLGEPLALAVSLEETGKITGNKAKVLLGKALDEATGKLDNNKS 646
QY 337 RGMEPLPERRL 347
DB 647 -----PSRKV 651

Tue Mar 8 09:44:21 2005

us-09-619-032a-4.ral

Page 8

Search completed: March 7, 2005, 21:45:20
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 21:42:11 ; Search time 132 Seconds

(Without alignments)
907.366 Million cell updates/sec

Title: US-09-619-032a-4

Sequence: 1 LRAVFGNLOVARIPIKSEI.....RRLDAFRAIYDNGENGP 364

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 329044822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1870	99.5	364	9 US-09-886-400-4	Sequence 4, Appl1
2	1870	99.5	364	13 US-10-112-357-4	Sequence 4, Appl1
3	1870	99.5	364	13 US-10-114-403-4	Sequence 4, Appl1
4	1870	99.5	364	13 US-10-116-406-4	Sequence 4, Appl1
5	1870	99.5	364	13 US-10-112-331-4	Sequence 4, Appl1
6	1870	99.5	364	13 US-10-112-377-4	Sequence 4, Appl1
7	1870	99.5	364	13 US-10-116-581-4	Sequence 4, Appl1
8	1870	99.5	364	13 US-10-112-442-4	Sequence 4, Appl1
9	1870	99.5	364	13 US-10-112-418-4	Sequence 4, Appl1
10	1870	99.5	364	13 US-10-114-083-4	Sequence 4, Appl1
11	111.5	5.9	890	15 US-10-282-122A-53281	Sequence 53281, A
12	109	5.8	312	15 US-10-369-493-175517	Sequence 175517, A
13	105.5	5.6	573	15 US-10-424-599-175517	Sequence 175517, A

14	105.5	5.6	730	15 US-10-425-114-55544	Sequence 55544, A
15	103	5.5	1073	9 US-09-815-242-12361	Sequence 12361, A
16	103	5.5	1147	9 US-09-815-242-5468	Sequence 5468, Ap
17	103	5.5	1150	15 US-10-282-122A-44391	Sequence 44391, A
18	102.5	5.5	478	15 US-10-369-493-4506	Sequence 4506, Ap
19	102.5	5.5	478	15 US-10-369-493-7264	Sequence 7264, Ap
20	101	5.4	441	16 US-10-437-963-122721	Sequence 122721, A
21	100.5	5.3	314	15 US-10-369-493-17385	Sequence 17385, A
22	98.5	5.2	813	15 US-10-282-122A-47208	Sequence 47208, A
23	98	5.2	510	15 US-10-369-493-56	Sequence 56, Appl1
24	97	5.2	652	15 US-10-282-122A-63055	Sequence 63055, A
25	97	5.2	1878	15 US-10-607-631-20	Sequence 20, Appl1
26	96.5	5.1	798	15 US-10-425-114-37809	Sequence 37809, A
27	96.5	5.1	1085	15 US-10-282-122A-58833	Sequence 58833, A
28	96.5	5.1	3432	15 US-10-282-122A-69849	Sequence 69849, A
29	96	5.1	530	15 US-10-369-493-10268	Sequence 10268, A
30	96	5.1	530	15 US-10-282-122A-77165	Sequence 77165, A
31	96	5.1	698	16 US-10-437-963-122301	Sequence 122301, A
32	95.5	5.1	875	15 US-10-282-122A-52807	Sequence 52807, A
33	95	5.1	535	15 US-10-282-122A-55460	Sequence 55460, A
34	95	5.1	2094	16 US-10-437-963-161607	Sequence 161607, A
35	94.5	5.0	749	15 US-10-369-493-9015	Sequence 9015, Ap
36	94.5	5.0	1822	15 US-10-432-443-39	Sequence 39, Appl1
37	94	5.0	274	15 US-10-282-122A-54582	Sequence 54582, A
38	94	5.0	664	14 US-10-032-585-7195	Sequence 7195, Ap
39	93.5	5.0	993	15 US-10-369-493-20014	Sequence 20014, A
40	93.5	5.0	1165	15 US-10-282-122A-71768	Sequence 71768, A
41	93.5	5.0	1320	14 US-10-161-051-167	Sequence 167, Ap
42	93.5	5.0	2049	16 US-10-437-963-197248	Sequence 197248, A
43	93.5	5.0	2111	16 US-10-437-963-186073	Sequence 186073, A
44	93.5	5.0	2498	16 US-10-437-963-186071	Sequence 186071, A
45	93	4.9	235	17 US-10-495-918-152	Sequence 152, Ap

ALIGNMENTS

RESULT 1
US-09-886-400-4
Sequence 4, Application US/09886400
Patent No. US20020045226A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: Murphy, Dennis
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
FILE REFERENCE: DIVER1120-4
CURRENT APPLICATION NUMBER: US/09/886,400
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 09/619,032
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/407,806
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 08/613,220
PRIOR FILING DATE: 1996-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 364
TYPE: PRT
ORGANISM: Thermococcus alcaliphilus
US-09-886-400-4

Query Match 99.5%, Score 1870, DB 9, Length 364,
Best Local Similarity 99.5%, Pred. No. 4.9e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRAVFGNLOVARIPIKSEIPIYETLILKEBIPFGNITGYTLKPIPKDII 60
DB 1 LRAVFGNLOVARIPIKSEIPIYETLILKEBIPFGNITGYTLKPIPKDII 60
QY 61 LVKGSISDLIEITGSTYTHAILPLPLPSREAQVODREVKELFEVSPGFMLPELAY 120

Db 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKEELFELSPKGFMLPELAY 120
Qy 121 DPLIPALIKONGYEYFADGEAMLFSAHLSAIPKIPKLYPHILIKQREKRFYISYLLG 180
Db 121 DPLIPALIKONGYEYFADGEAMLFSAHLSAIPKIPKLYPHILIKQREKRFYISYLLG 180
Qy 181 LRELRAIKLVFEGKVTLLKAVKOIEAVPVWVAVTAVMLGIGRLPLMNPCKVASWIEDKO 240
Db 181 LRELRAIKLVFEGKVTLLKAVKOIEAVPVWVAVTAVMLGIGRLPLMNPCKVASWIEDKO 240
Qy 241 NILLYGTDFIEFGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
Db 241 NILLYGTDFIEFGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
Qy 301 DKSRLIWRDEGNARLNLMSYNNRGELAFLAENSDARGWEPPLPERRLDAFRALYNDMRGE 360
Db 301 DKSRLIWRDEGNARLNLMSYNNRGELAFLAENSDARGWEPPLPERRLDAFRALYNDMRGE 360
Qy 361 NGEF 364
Db 361 NGEF 364

RESULT 2
US-10-112-357-4
; Sequence 4, Application US/10112357
; Publication No. US20020115099A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE THEREOF
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/112,357
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-357-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 4.9e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LRALVPHGNLQYAEIPKSEIPKVIKAYIPVETLLIKEBIPFGINTGYTLKFLPKDIIID 60
Db 1 LRALVPHGNLQYAEIPKSEIPKVIKAYIPVETLLIKEBIPFGINTGYTLKFLPKDIIID 60
Qy 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKEELFELSPKGFMLPELAY 120
Db 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKEELFELSPKGFMLPELAY 120
Qy 121 DPLIPALIKONGYEYFADGEAMLFSAHLSAIPKIPKLYPHILIKQREKRFYISYLLG 180
Db 121 DPLIPALIKONGYEYFADGEAMLFSAHLSAIPKIPKLYPHILIKQREKRFYISYLLG 180
Qy 181 LRELRAIKLVFEGKVTLLKAVKOIEAVPVWVAVTAVMLGIGRLPLMNPCKVASWIEDKO 240
Db 181 LRELRAIKLVFEGKVTLLKAVKOIEAVPVWVAVTAVMLGIGRLPLMNPCKVASWIEDKO 240
Qy 241 NILLYGTDFIEFGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
Db 241 NILLYGTDFIEFGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300

Db 241 NILLYGTDFIEFGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
Qy 301 DKSRLIWRDEGNARLNLMSYNNRGELAFLAENSDARGWEPPLPERRLDAFRALYNDMRGE 360
Db 301 DKSRLIWRDEGNARLNLMSYNNRGELAFLAENSDARGWEPPLPERRLDAFRALYNDMRGE 360
Qy 361 NGEF 364
Db 361 NGEF 364

RESULT 3
US-10-114-403-4
; Sequence 4, Application US/10114403
; Publication No. US20020115100A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE THEREOF
; FILE REFERENCE: DIVER1120-4
; CURRENT APPLICATION NUMBER: US/10/114,403
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-114-403-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 4.9e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LRALVPHGNLQYAEIPKSEIPKVIKAYIPVETLLIKEBIPFGINTGYTLKFLPKDIIID 60
Db 1 LRALVPHGNLQYAEIPKSEIPKVIKAYIPVETLLIKEBIPFGINTGYTLKFLPKDIIID 60
Qy 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKEELFELSPKGFMLPELAY 120
Db 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVDRREVKEELFELSPKGFMLPELAY 120
Qy 121 DPLIPALIKONGYEYFADGEAMLFSAHLSAIPKIPKLYPHILIKQREKRFYISYLLG 180
Db 121 DPLIPALIKONGYEYFADGEAMLFSAHLSAIPKIPKLYPHILIKQREKRFYISYLLG 180
Qy 181 LRELRAIKLVFEGKVTLLKAVKOIEAVPVWVAVTAVMLGIGRLPLMNPCKVASWIEDKO 240
Db 181 LRELRAIKLVFEGKVTLLKAVKOIEAVPVWVAVTAVMLGIGRLPLMNPCKVASWIEDKO 240
Qy 241 NILLYGTDFIEFGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
Db 241 NILLYGTDFIEFGYRDIAGYRMSVEGLLEVIDELNSELCLPSELKHSGRELYLRTSSWAP 300
Qy 301 DKSRLIWRDEGNARLNLMSYNNRGELAFLAENSDARGWEPPLPERRLDAFRALYNDMRGE 360
Db 301 DKSRLIWRDEGNARLNLMSYNNRGELAFLAENSDARGWEPPLPERRLDAFRALYNDMRGE 360
Qy 361 NGEF 364
Db 361 NGEF 364

RESULT 4
US-10-116-606-4
Sequence 4, Application US/10116606
Publication No. US20020119515A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: Murphy, Dennis
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE
FILE REFERENCE: DIVER1120-4
CURRENT APPLICATION NUMBER: US/10/116,606
PRIOR FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US/09/886,400
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 09/619,032
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/407,806
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 08/613,220
PRIOR FILING DATE: 1996-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 364
TYPE: PRT
ORGANISM: Thermococcus alcaliphilus
US-10-116-606-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 4.9e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEIPFGNLITGYTLKFLPKDIIID 60
DB 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEIPFGNLITGYTLKFLPKDIIID 60
QY 61 LVKGIASDLLEIIGTSTYTHAILPLPLSRVBAOVORREVEKEELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLLEIIGTSTYTHAILPLPLSRVBAOVORREVEKEELFEVSPKGFMLPELAY 120
QY 121 DPIIPALIKDNGEYELFADGEMLFSAHLNSAIKIPKLYPHLIIKAQREKFRYISYLLG 180
DB 121 DPIIPALIKDNGEYELFADGEMLFSAHLNSAIKIPKLYPHLIIKAQREKFRYISYLLG 180
QY 181 LRELKAIKLVFEGKVTLKAVKDIEAVPVWVAVNTAVMLGIGRLPLNPKKVASWIEDKD 240
DB 181 LRELKAIKLVFEGKVTLKAVKDIEAVPVWVAVNTAVMLGIGRLPLNPKKVASWIEDKD 240
QY 241 NILYGTDIEFTIGYRDIAGYMSVEGLEVIDELNSELCPLSELKHSGRELYLRTSSWAP 300
DB 241 NILYGTDIEFTIGYRDIAGYMSVEGLEVIDELNSELCPLSELKHSGRELYLRTSSWAP 300
QY 301 DKSLRIMREDEGNARLNLSTYMRGELAFLENSDARGMEPLPERRLDAFPAIYNDMRGE 360
DB 301 DKSLRIMREDEGNARLNLSTYMRGELAFLENSDARGMEPLPERRLDAFPAIYNDMRGE 360
QY 361 NGEF 364
DB 361 NGEF 364

RESULT 5
US-10-112-331-4
Sequence 4, Application US/10112331
Publication No. US20020119550A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: Murphy, Dennis
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE
FILE REFERENCE: DIVER1120-4
CURRENT APPLICATION NUMBER: US/10/112,331
CURRENT FILING DATE: 2002-03-29

PRIOR APPLICATION NUMBER: US/09/886,400
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 09/619,032
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/407,806
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 08/613,220
PRIOR FILING DATE: 1996-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 364
TYPE: PRT
ORGANISM: Thermococcus alcaliphilus
US-10-112-331-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 4.9e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEIPFGNLITGYTLKFLPKDIIID 60
DB 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEIPFGNLITGYTLKFLPKDIIID 60
QY 61 LVKGIASDLLEIIGTSTYTHAILPLPLSRVBAOVORREVEKEELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLLEIIGTSTYTHAILPLPLSRVBAOVORREVEKEELFEVSPKGFMLPELAY 120
QY 121 DPIIPALIKDNGEYELFADGEMLFSAHLNSAIKIPKLYPHLIIKAQREKFRYISYLLG 180
DB 121 DPIIPALIKDNGEYELFADGEMLFSAHLNSAIKIPKLYPHLIIKAQREKFRYISYLLG 180
QY 181 LRELKAIKLVFEGKVTLKAVKDIEAVPVWVAVNTAVMLGIGRLPLNPKKVASWIEDKD 240
DB 181 LRELKAIKLVFEGKVTLKAVKDIEAVPVWVAVNTAVMLGIGRLPLNPKKVASWIEDKD 240
QY 241 NILYGTDIEFTIGYRDIAGYMSVEGLEVIDELNSELCPLSELKHSGRELYLRTSSWAP 300
DB 241 NILYGTDIEFTIGYRDIAGYMSVEGLEVIDELNSELCPLSELKHSGRELYLRTSSWAP 300
QY 301 DKSLRIMREDEGNARLNLSTYMRGELAFLENSDARGMEPLPERRLDAFPAIYNDMRGE 360
DB 301 DKSLRIMREDEGNARLNLSTYMRGELAFLENSDARGMEPLPERRLDAFPAIYNDMRGE 360
QY 361 NGEF 364
DB 361 NGEF 364

RESULT 6
US-10-112-377-4
Sequence 4, Application US/10112377
Publication No. US20020120108A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: Murphy, Dennis
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE
FILE REFERENCE: DIVER1120-4
CURRENT APPLICATION NUMBER: US/10/112,377
CURRENT FILING DATE: 2002-03-29
PRIOR APPLICATION NUMBER: 09/886,400
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 09/619,032
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/407,806
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 08/613,220
PRIOR FILING DATE: 1996-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 364

TYPE: PRT
ORGANISM: Thermococcus alcaliphilus
US-10-112-377-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 4.9e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLVHGNLQVAEIPKSEIPKVIKAYIPVITLKEEIPGLNTGTTLKFLPDIID 60
DB 1 LRLVHGNLQVAEIPKSEIPKVIKAYIPVITLKEEIPGLNTGTTLKFLPDIID 60
QY 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVORDREVKEELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVORDREVKEELFEVSPKGFMLPELAY 120
QY 121 DPLIPALKONGEYELFADGEAMLFSAHLSAIIKPIKPLYPHLIKAKQREKRFYISYLLG 180
DB 121 DPLIPALKONGEYELFADGEAMLFSAHLSAIIKPIKPLYPHLIKAKQREKRFYISYLLG 180
QY 181 LRELRAIKLVEGKVTLLKAVKDIEAVPVVAVNTAVMTGIGRLPLMNPCKVASWIEDKD 240
DB 181 LRELRAIKLVEGKVTLLKAVKDIEAVPVVAVNTAVMTGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILVGTDFEIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRLEYLRTSSMAP 300
DB 241 NILVGTDFEIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRLEYLRTSSMAP 300
QY 301 DKSRLIWRDEGNARLNMLSYNNRGELAPLAENSDARGWEPLPERLDAFRAIYNDWRGE 360
DB 301 DKSRLIWRDEGNARLNMLSYNNRGELAPLAENSDARGWEPLPERLDAFRAIYNDWRGE 360
QY 361 NGEIP 364
DB 361 NGEIP 364

RESULT 7
US-10-116-581-4

Sequence 4, Application US/10116581
Publication No. US20020137116A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: Murphy, Dennis
APPLICANT: Ried, John
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: DIVER1120-4
CURRENT APPLICATION NUMBER: US/10/116,581
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 09/886,400
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 09/619,032
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/407,806
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 08/613,220
PRIOR FILING DATE: 1996-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 364
TYPE: PRT
ORGANISM: Thermococcus alcaliphilus
US-10-116-581-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 4.9e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLVHGNLQVAEIPKSEIPKVIKAYIPVITLKEEIPGLNTGTTLKFLPDIID 60
DB 1 LRLVHGNLQVAEIPKSEIPKVIKAYIPVITLKEEIPGLNTGTTLKFLPDIID 60

QY 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVORDREVKEELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVORDREVKEELFEVSPKGFMLPELAY 120
QY 121 DPLIPALKONGEYELFADGEAMLFSAHLSAIIKPIKPLYPHLIKAKQREKRFYISYLLG 180
DB 121 DPLIPALKONGEYELFADGEAMLFSAHLSAIIKPIKPLYPHLIKAKQREKRFYISYLLG 180
QY 181 LRELRAIKLVEGKVTLLKAVKDIEAVPVVAVNTAVMTGIGRLPLMNPCKVASWIEDKD 240
DB 181 LRELRAIKLVEGKVTLLKAVKDIEAVPVVAVNTAVMTGIGRLPLMNPCKVASWIEDKD 240
QY 241 NILVGTDFEIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRLEYLRTSSMAP 300
DB 241 NILVGTDFEIGYRDIAGRMSVEGLLEVIDELNSELCLPSELKHSGRLEYLRTSSMAP 300
QY 301 DKSRLIWRDEGNARLNMLSYNNRGELAPLAENSDARGWEPLPERLDAFRAIYNDWRGE 360
DB 301 DKSRLIWRDEGNARLNMLSYNNRGELAPLAENSDARGWEPLPERLDAFRAIYNDWRGE 360
QY 361 NGEIP 364
DB 361 NGEIP 364

RESULT 8
US-10-112-442-4

Sequence 4, Application US/10112442
Publication No. US20020150997A1
GENERAL INFORMATION:
APPLICANT: DIVERSA CORPORATION
APPLICANT: Murphy, Dennis
APPLICANT: Ried, John
TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND METHODS OF USE TH
FILE REFERENCE: DIVER1120-4
CURRENT APPLICATION NUMBER: US/10/112,442
PRIOR FILING DATE: 2002-03-29
PRIOR APPLICATION NUMBER: 09/886,400
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 09/619,032
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/407,806
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 08/613,220
PRIOR FILING DATE: 1996-03-08
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 364
TYPE: PRT
ORGANISM: Thermococcus alcaliphilus
US-10-112-442-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 4.9e-165;
Matches 352; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLVHGNLQVAEIPKSEIPKVIKAYIPVITLKEEIPGLNTGTTLKFLPDIID 60
DB 1 LRLVHGNLQVAEIPKSEIPKVIKAYIPVITLKEEIPGLNTGTTLKFLPDIID 60
QY 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVORDREVKEELFEVSPKGFMLPELAY 120
DB 61 LVKGIASDLIEIIGTSYTHAILPLPLSRVEAQQVORDREVKEELFEVSPKGFMLPELAY 120
QY 121 DPLIPALKONGEYELFADGEAMLFSAHLSAIIKPIKPLYPHLIKAKQREKRFYISYLLG 180
DB 121 DPLIPALKONGEYELFADGEAMLFSAHLSAIIKPIKPLYPHLIKAKQREKRFYISYLLG 180
QY 181 LRELRAIKLVEGKVTLLKAVKDIEAVPVVAVNTAVMTGIGRLPLMNPCKVASWIEDKD 240
DB 181 LRELRAIKLVEGKVTLLKAVKDIEAVPVVAVNTAVMTGIGRLPLMNPCKVASWIEDKD 240

```
Qy 241 NILYGTDIEFGYDIAGYRMSVGLLEVIDELNSELCPLSELKHSRELYLRTSSWAP 300
Db 241 NILYGTDIEFGYDIAGYRMSVGLLEVIDELNSELCPLSELKHSRELYLRTSSWAP 300
Qy 301 DKSRLRWEDEGNAALNMLSYNMRGELAFLENSDARGMEPLPERRLDAFRAIYNDWGE 360
Db 301 DKSRLRWEDEGNAALNMLSYNMRGELAFLENSDARGMEPLPERRLDAFRAIYNDWGE 360
Qy 361 NGEF 364
Db 361 NGEF 364

RESULT 9
US-10-112-418-4
; Sequence 4, Application US/10112418
; Publication No. US20020155486A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
; FILE REFERENCE: DIVERS1120-4
; CURRENT APPLICATION NUMBER: US/10/112,418
; PRIOR FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-112-418-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 4,9e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEIPGKLNITGYTLKFLPKDIID 60
Db 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEIPGKLNITGYTLKFLPKDIID 60
Qy 61 LVKGGIASDLIEIGTSTYTHAILPLPLSRVEAQVQDRREVEKELFEVSPKGFMLPELAY 120
Db 61 LVKGGIASDLIEIGTSTYTHAILPLPLSRVEAQVQDRREVEKELFEVSPKGFMLPELAY 120
Qy 121 DPLIPALIKDNGEYELFADGEMLPSAHLNSAIKIKPLYPHLIAQREKFRYISYLLG 180
Db 121 DPLIPALIKDNGEYELFADGEMLPSAHLNSAIKIKPLYPHLIAQREKFRYISYLLG 180
Qy 181 LRELKAIKLVPEGVTLKAVKDI EAVPVWVA VNTAVNLGIGRLPLMMPKVAASVIEDKD 240
Db 181 LRELKAIKLVPEGVTLKAVKDI EAVPVWVA VNTAVNLGIGRLPLMMPKVAASVIEDKD 240
Qy 241 NILYGTDIEFGYDIAGYRMSVGLLEVIDELNSELCPLSELKHSRELYLRTSSWAP 300
Db 241 NILYGTDIEFGYDIAGYRMSVGLLEVIDELNSELCPLSELKHSRELYLRTSSWAP 300
Qy 301 DKSRLRWEDEGNAALNMLSYNMRGELAFLENSDARGMEPLPERRLDAFRAIYNDWGE 360
Db 301 DKSRLRWEDEGNAALNMLSYNMRGELAFLENSDARGMEPLPERRLDAFRAIYNDWGE 360
Qy 361 NGEF 364
Db 361 NGEF 364
```

```
Db 361 NGEF 364

RESULT 10
US-10-114-083-4
; Sequence 4, Application US/10114083
; Publication No. US20020160464A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: Murphy, Dennis
; APPLICANT: Ried, John
; TITLE OF INVENTION: ENZYMES HAVING ALPHA-GALACTOSIDASE ACTIVITY AND
; FILE REFERENCE: DIVERS1120-4
; CURRENT APPLICATION NUMBER: US/10/114,083
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/886,400
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 09/619,032
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/407,806
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 08/613,220
; PRIOR FILING DATE: 1996-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Thermococcus alcaliphilus
US-10-114-083-4

Query Match 99.5%; Score 1870; DB 13; Length 364;
Best Local Similarity 99.5%; Pred. No. 4,9e-165;
Matches 362; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEIPGKLNITGYTLKFLPKDIID 60
Db 1 LRALVFHGNLQYAEIPKSEIPKVIKAYIPVETLKEIPGKLNITGYTLKFLPKDIID 60
Qy 61 LVKGGIASDLIEIGTSTYTHAILPLPLSRVEAQVQDRREVEKELFEVSPKGFMLPELAY 120
Db 61 LVKGGIASDLIEIGTSTYTHAILPLPLSRVEAQVQDRREVEKELFEVSPKGFMLPELAY 120
Qy 121 DPLIPALIKDNGEYELFADGEMLPSAHLNSAIKIKPLYPHLIAQREKFRYISYLLG 180
Db 121 DPLIPALIKDNGEYELFADGEMLPSAHLNSAIKIKPLYPHLIAQREKFRYISYLLG 180
Qy 181 LRELKAIKLVPEGVTLKAVKDI EAVPVWVA VNTAVNLGIGRLPLMMPKVAASVIEDKD 240
Db 181 LRELKAIKLVPEGVTLKAVKDI EAVPVWVA VNTAVNLGIGRLPLMMPKVAASVIEDKD 240
Qy 241 NILYGTDIEFGYDIAGYRMSVGLLEVIDELNSELCPLSELKHSRELYLRTSSWAP 300
Db 241 NILYGTDIEFGYDIAGYRMSVGLLEVIDELNSELCPLSELKHSRELYLRTSSWAP 300
Qy 301 DKSRLRWEDEGNAALNMLSYNMRGELAFLENSDARGMEPLPERRLDAFRAIYNDWGE 360
Db 301 DKSRLRWEDEGNAALNMLSYNMRGELAFLENSDARGMEPLPERRLDAFRAIYNDWGE 360
Qy 361 NGEF 364
Db 361 NGEF 364

RESULT 11
US-10-282-53281
; Sequence 53281, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
```

```

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITTA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53281
; LENGTH: 890
; TYPE: PRT
; ORGANISM: Clostridium difficile
; US-10-282-122A-53281

Query Match
Best Local Similarity 23.6%; Score 111.5; DB 15; Length 890;
Matches 72; Conservative 48; Mismatches 98; Indels 87; Gaps 16;
```

```

QY 53 FLPKDIIIDLVKGIASDLIEIIGTSTHAIPLPLSRVEAOVORDREVEKEELFEVSPKG 112
DB 562 FQP-DLIIVAGGGSAMDAGKIMWVMEHPVEDFODLA-----MRFMDIKRKY-VFPK- 612
QY 113 FWLPBELAYDPIIP-----AIL--KONGYEYLPADGEAMLFSAHLNSAIKPIKP 158
DB 613 --NGBEAYFPAITTSAGTSEVTPFAVITTDQDSGVKPLDVELMPNMAIIDDMMEMP 670
QY 159 LYPHL-----IKAQREKRFRIYSYL-----LGLBELKRAIKLVFE-----GKVTL 198
DB 671 --PRLTAASGVDLTLHALLEVAVSMLRTEPADGLAL-----QAKGIIPEYLRAYKNGNDK 724
QY 199 KAVKDIEAVPVVAVNTA-----VNLGIGRLPLAN-----P 229
DB 725 EAREKXAMASTWAGSFANALFICISLAHKLGAFFHVQGVANALINIVIKENCAP 784
QY 230 KKVASMIEDK--DNILLYGTDIRFIGYRDYAG--RMSVGLGLEVIDELNSEICLPSELK 285
DB 785 NKKGAISQRYKPCIORYA--EFASPAKIGSTDEKDNLIKALDELAKVGLPKTK 841
QY 286 HSGRE 290
DB 842 EAGVE 846
```

RESULT 12
US-10-369-493-23237

```

; Sequence 23237, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23237
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Bacillus subtilis
; US-10-369-493-23237

Query Match
Best Local Similarity 23.2%; Score 109; DB 15; Length 312;
Matches 44; Conservative 32; Mismatches 58; Indels 56; Gaps 9;
```

```

QY 10 LCVAAEIPKSEIPK-----VIEKAYIVIEITLKEEIPFGINIGYTLKFLPKDIIIDL 61
DB 130 MVAAYKESGFPERYIGSGVLDTR---FRTFAVE-----LNLG-----VKDYTG 175
QY 62 VVGGIASDLIEIIGTSTHAI--LPLPLSRVEAOVORDREVEKEELFEVSPKG--FWLPE 117
DB 176 VLGHGDDWVPLRVSYAGIPLFETLIPKRIIDAIYERRKGGGELVNLGNASAYAPA 235
QY 118 LAYDPIIPALIKDN-----GYEYLF-----ADGEAMLF-----A 147
DB 236 FSLTEWVEAILKQORVPLPTIAYLGEYGEIGYLVPTIVGGNGLEQIIEELTYERA 295
QY 148 HLNSAIKPIK 157
DB 296 QLNKSYESVK 305
```

```

RESULT 13
US-10-424-599-175517
; Sequence 175517, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175517
; LENGTH: 573
; TYPE: PRT
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12950C.1.pep
; US-10-424-599-175517
```

Query Match
Best Local Similarity 21.3%; Score 105.5; DB 15; Length 573;
Matches 77; Conservative 57; Mismatches 97; Indels 131; Gaps 19;

QY 6 FHGNL-QYAEIPKS-----EIPKVIKAYIPV-----IETL 35
DB 290 FLGSLCKLGGVPSLGLSGVBPVIVSPFSPYCWCPGISTCPSIAAVTQSPNSSIETL 349

Query Match	5.6%	Score 105.5	DB 15	Length 730
Best Local Similarity	21.3%	Pred No. 1.3		
Matches	77	Conservative	57	Mismatches 97
				Indels 13
				Gaps 19

QY	6	FHGNL-QYAEIPKS-----EIPKVIKAYIPV-----	-----IETL	35
Db	447	FLGISIGKLGVPSSLSGGEVPIVSPFSPFYWCPCPGISRCPSIAAVTQSPNSIETL	-----	506
QY	36	IKESIFPEPLANTGTYLTKFLPKDIIDLVKGLASDLIEITGTSYTHALPLPLRSKEYAO	-----	95
Db	507	---PFPSSASLAPLPLS-----VNL-----DVPQPLGISMDDP--PFLP--	-----DPLV	545
QY	96	QRDRVEKKELEFEVSPKGFMLPELAPDPI--IPAI-LKONGVEYELPADGEAMLFSALHNSA	-----	152
Db	546	RMSLPTSQOIPFTL-----LMCDPIYHVNVIVDVCSSGGQYLVSAGPANSPS-----	-----	592
QY	153	IKPIKPLVPHLIKAGREKRFRIYSYLLGLRELKAIKLVFEG-----KVTIKAYKD	-----	203
Db	593	---IPPLHPIIVKPIIPESDAVV-----KGARETLRLLSGSSGQGNQOOWMDLPAL--	-----	641
QY	204	IEAVPVVAVNTAVWLIGIRLPLMPKPKVYASIEEDKNLILVGTIEIFIGRDI-----	-----	257

```

Db      642 -----LTPD-----ENQNNILVAGSGLYGTGRDIAMANS 673
QY      258 --AGRMSVEGLLEVIDELNSGLCLPSLEKSHGRELTYLTSSWADPKSLRTWRDEGNAR 315
Db      674 IAAAGLVSLISGVSKVDGYSSEL.C-----ENYGNLEAVKNSN---DSGGAGLDDGGSS 725
QY      316 LN 317
Db      726 LD 727

RESULT 15
US-09-815-242-12361
; Sequence 12361, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykend, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12361
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12361

Query Match      5.5%: Score 103; DB 9; Length 1073;
Best Local Similarity 21.2%: Pred. No. 4;
Matches 58; Conservative 48; Mismatches 90; Indels 78; Gaps 15

QY      107 EVSPKGFMLPELAYDPILAIKONGVEYELPADGEAMLFSAHLSAIPKIPYPLHIKA 166
Db      309 EISP-----YDSSL---LVKSTHAISFKQAEEKVVRSLREMRIRNGVKNIPFLNV 357
QY      167 QREKRF-----RYSISLILGLRELKRAI-----KLVEGKYTLKAVKDIE----- 205
Db      358 MNKKKFTSGDYTKRIETPELPFDIQLDGTGKLEIYGVNTINGPENVKRPKPDYEL 417
QY      206 -AVPVMVAVNTAVMVLGIGR-LPLNPPKVASMIEDKNIILYGTDIEF-IGYRDIAGYRM 262
Db      418 ASIPVSSSKIASSGTKQLDDEYGPKVAAEWKKDDVLL-IDTTRDAHQSLATRV 475
QY      263 SVEGLEVIDELNSBELCLPSELKSHGRELTYLTSSWAPDK-SLRTWRDEGNARLMLSY 321
Db      476 RTKGMINIAS-----KTDVFPDGGFSLWEM---GGATPD-VAY 509
QY      322 NMKGELAFIANSDARGCEPLPERRLDAFRAIYN 355

```

Db 510 N-----FLKEND---WERLRLR---KAIPN 529

Search completed: March 7, 2005, 21:56:19
Job time : 134 secs
